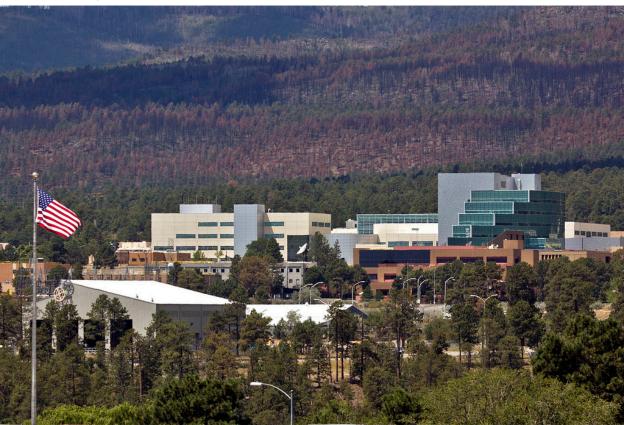# EDGE bioinformatics brings genomics to everyone

November 29, 2016

New platform enables rapid genomic-sequence data analysis

LOS ALAMOS, N.M., Nov. 29, 2016 – A new bioinformatics platform called Empowering the Development of Genomics Expertise (EDGE) will help democratize the genomics revolution by allowing users with limited bioinformatics expertise to quickly analyze and interpret genomic sequence data. Researchers at Los Alamos National Laboratory and their collaborators at the Naval Medical Research Center developed EDGE, which is described in a paper recently published in *Nucleic Acids Research*.

"We realized that while next-generation sequencing instruments are becoming more widespread and more accessible to the average biologist or physician, the bioinformatics tools required to process and analyze the data were not as user-friendly or accessible," said Patrick Chain, of Los Alamos' Biosecurity and Public Health group

and EDGE team lead. "Given the large number of applications where sequencing is now used, a robust bioinformatics platform that encapsulates a broad array of algorithms is required to help address questions a researcher may have. We sought to develop a web-based environment where non-bioinformatics experts could easily select what pipelines they need and rapidly obtain results and interact with their data."

Stopping the spread of disease—from naturally occurring or manmade threats—requires an in-depth understanding of pathogens and how they work. To this end, the ability to characterize organisms through accurately and rapidly comparing genomic data is an important part of Los Alamos' national security mission.

Technology advancements have fueled the development of new sequencing applications and will flood current databases with raw data. A number of factors limit the use of these data, including the large number of associated software and hardware dependencies and the detailed expertise required to perform this analysis. To address these issues, Chain and his team have developed an intuitive web-based environment with a wide assortment of integrated and pioneering bioinformatics tools in pre-configured workflows, all of which can be readily applied to isolate genome sequencing projects or metagenomics projects.

EDGE is a user-friendly and open-source platform that integrates hundreds of cutting-edge tools and helps reduce data analysis times from days or weeks to minutes or hours. The workflows in EDGE, along with its ease of use, provide novice next-generation sequencing users with the ability to perform many complex analyses with only a few mouse clicks. This bioinformatics platform is described as an initial attempt at empowering the development of genomics expertise, as its name suggests, for a wide range of applications in microbial research.

EDGE has already helped streamline data analysis for groups in Thailand, Georgia, Peru, South Korea, Gabon, Uganda, Egypt and Cambodia, as well as within several government laboratories in the United States.

The paper "[Enabling the democratization of the genomics revolution with a fully integrated web-based bioinformatics platform](#)" was published in Nucleic Acids Research in partnership with the Defense Threat Reduction Agency, the Naval Medical Research Center-Frederick and the Henry M. Jackson Foundation.

About Chain

Patrick Chain earned his master's of science in microbial genomics from McMaster University and his doctoral degree in molecular microbiology and molecular genetics at Michigan State University. He is currently leading the Bioinformatics and Analytics Team and the Metagenomics Program within the Biosecurity and Public Health group at Los Alamos National Laboratory.

His background is in microbial ecology, evolution, genomics and bioinformatics, having spent the past 20 years using genomics to study various microbial systems, including the human microbiome, other environmental metagenomic communities, various isolate microbes or single cells, including bacterial and viral pathogens as well as fungal, algal, plant and animal systems. He currently leads a team of researchers whose charge is to devise novel methods, algorithms and strategies for the biological interpretation of massively parallel sequencing data.

Operated by Los Alamos National Security, LLC for the Department of Energy's NNSA